Investigation of Complex Command, Control and Communications Decisionmaking under Sustained Operations

Linda R. Elliott
Veridian Engineering
Brooks AFB, TX
(210) 536-8137
linda.elliott@brooks.af.mil

James C. Miller
Chronobiology and Sleep Laboratory
Air Force Research Laboratory
Brooks AFB, TX
(210) 536-6371
jcmiller@brooks.af.mil

Christopher Barnes
Warfighter Training Research Division
Air Force Research Laboratory
Brooks AFB, TX
(210) 536-2177
christopher.barnes@brooks.af.mil

Mathieu Dalrymple
Veridian Engineering
Brooks AFB, TX
(210) 534-1134
mathieu.dalrymple@brooks.af.mil

Leroy Brown
Warfighter Training Research Division
Air Force Research Laboratory
Brooks AFB, TX
(210) 536-2144
leroy.brown@brooks.af.mil

Jeff Whitmore
Veridian Engineering
Brooks AFB, TX 78235
(210) 536-2177
jeff.whitmore@brooks.af.mil

Joe Fischer
Veridian Engineering
Brooks AFB, TX 78235
(210) 536-1134
joe.fischer@brooks.af.mil

Rebecca Cardenas
NTI, Inc
Brooks AFB, TX
(210) 536-2177
rebecca.cardenas@brooks.af.mil

Abstract: In this paper we describe plans and initial progress in baseline investigations of fatigue on team performance in complex and operationally relevant task environments. Preliminary data collection used a PC-based analogue of command and control simulations. The platform was developed based on cognitive and functional analysis of C3 mission, tactics, team member roles, and role interdependencies. Tactical scenarios were developed to capture core team coordination, decision-making and problem-solving task demands. Issues regarding measures and scenario development are identified and discussed. Preliminary findings, indicating increased resistance to fatigue effects over time, are presented. Lessons learned are noted, along with plans for subsequent research.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **SEP 2002**	2. REPORT TYPE	3. DATES COVERED **00-00-2002 to 00-00-2002**

4. TITLE AND SUBTITLE **Investigation of Complex Command, Control and Communications Decisionmaking under Sustained Operations**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Veridian Engineering,Brooks AFB,TX,78235**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**13**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Introduction

Sustained operations are integral to command and control—combat missions require vigilance over time and adaptive performance under stress. Situations requiring close coordination and adaptive replanning are increasingly prevalent and challenging. Requirements for multi-service coordination are increasing, in maneuvers that are mobile, rapid, dynamic, and constantly evolving. Current examples include tactics such as battlefield interdiction and close air support in situations requiring rapid movement of troops and armament.

Advanced technology enables closer coordination and accuracy of long-range weapons; however, technology increases the demand for human performance--for close team and multi-team coordination, shared situation awareness across numerous and diverse units, and rapid replanning across units, in hostile, dynamic, time-critical, and long-duration situations. In fact, technology actually increases the role and demands of the human decision maker.

While advanced technology affords paradigm shifts in information technology, it cannot replace C3 decision makers or troops on the ground. Both are challenged to make tactical decisions under duress, often for long periods of time. Consider the following statement offered by an active-duty special forces member, describing the impact of new demands on existing tactics, and the fundamental role of fatigue in battle:

> "The Light forces on foot go into battle with 3 days of supplies. LRSU and SF might stretch this to a week or 2 using heavy rucksacks and compressed (dehydrated rations). The Heavy units in ground vehicles go in with 24 hours of fuel and 3 days of ammo/food but are expected thereafter to fight indefinately re: IDF in Palestinian territories today. A vehicle can carry bulk water, food, ammo, a human back cannot. ...since the Light forces after 3 days of walking around with heavy rucks need resupply but are in need of quality rest off the "line" (if there is a FLOT –Forward Line of Troops) re: the HBO mini-series "Band of Brothers" showing Paratroopers in a non-linear warfare situation is a good illustration. Once a FLOT was established, other units could replace the Airborne on the "line" to give them a rest. We haven't had this kind of manpower to have a second and third echelon to spell the first echelon since WWII. Since WWII there has been only one echelon of troops fighting, and they have to carry everything on their backs and recover their own dead and wounded. …However, Operation Anaconda in Afghanistan recently stretched on for over 1 week (was supposed to be 3 days) and the helicopter-delivered lightfighters were exhausted and the cordon around the enemy had to be released."

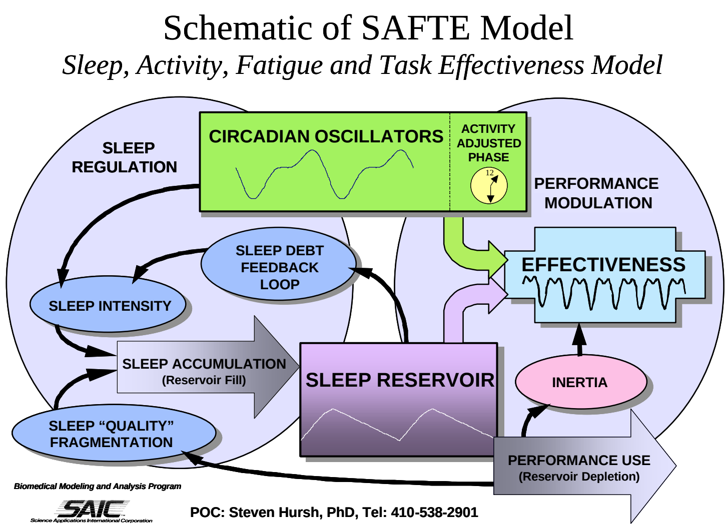
Of course, the fundamental importance of fatigue in battle has always been acknowledged:

"You must not needlessly fatigue the troops."
Napoleon Bonaparte[1]

The importance of fatigue is without doubt. Research has afforded much insight with regard to the impact (and measurement) of fatigue on various aspects of human performance. Advancements in theory and measurement have led to development of quantitative models for prediction of fatigue effects, and to fatigue countermeasures. The predictive model has been optimized to predict changes in cognitive performance in addition to alertness and incorporates

[1] To the Armée d'Italie in 1796.

features such as: a multi-oscillator circadian process, a circadian sleep propensity process, a sleep fragmentation process, and a circadian phase adjusting feature for time zone changes (Eddy &Hursh, 2000; Hursh, 1998).



However, if we turn our attention to effects of fatigue on team or mission performance, it is surprising to find very little experimental, controlled studies of the effect of fatigue on team performance. While effects of fatigue on basic cognitive functions have been investigated from a number of perspectives, there is less experimental data on effects of fatigue in more operational task contexts, and almost none with regard to operational team performance. In fact, very few field studies report effects on aspects of teamwork, even if performance was observed in isolated long-duration teams. In contrast, anecdotal evidence consistently report effects are quite detrimental.

The importance and multi-dimensional nature of teamwork in C3 operations is well known. While it is relatively simple to predict detrimental effects of fatigue, how do we become more specific? What sorts of error will occur, when, by whom, under what conditions? Will experience and expertise moderate effects, and if so, how? How will fatigue effect specific aspects of teamwork? Will there be a general compensatory effect such that teams could be less affected by fatigue—or will there be detrimental effect due to tunneling of attention and/or irritability?

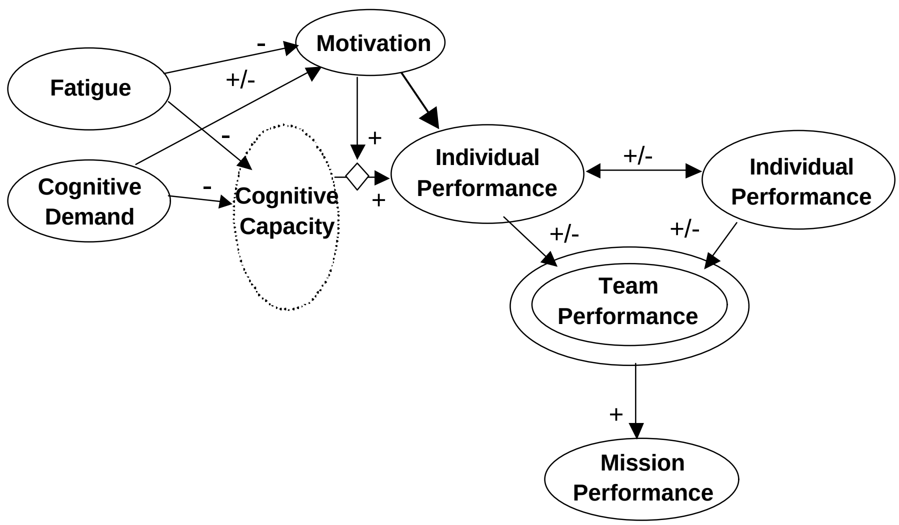
Questions are numerous. Therefore, attention and effort is turning to investigate and model effects of fatigue and chronobiology on command and control team decisionmaking. This paper describes overall plans, methodology, and metrics. We describe plans and initial progress in baseline investigations of fatigue on team cognitive function, decisionmaking and performance in complex and operationally relevant task environments.

Preliminary data collection will occur using PC-based analogues of command and control simulations. These platforms were developed based on cognitive and functional analysis of C3 mission, tactics, teammember roles, and role interdependencies. Tactical scenarios were developed to capture core team coordination, decision-making and problem-solving task demands. These platforms provide an advanced platform for research and/or training, with high experimental control, manipulation, and online performance monitoring capabilities. The advantages of these capabilities are increased experimental control, manipulation, and

operational relevance (Bowers, Salas, Prince, & Brannick, 1992; Cannon-Bowers, Burns, Salas, & Pruitt, 1998; Schiflett & Elliott, 2000).

Model

Our initial conceptual framework is represented below. Data collection and analyses will be conducted consistent with hypotheses generated by the framework with the aim of capturing core constructs and measures associated with effects of fatigue on team performance. As data accumulates, the model will be tested and refined.



The framework depicts expected relationships among several general constructs. Some of the constructs are very broad, such as team performance, and are more fully specified separately. Details of each category will be specified through planned research efforts.

First, the framework predicts that fatigue will reduce cognitive capacity of decision makers. Cognitive capacity has been described in several resource-allocation models ((Kanfer, 1991; Kanfer & Ackerman, 1989) as the degree to which attention and ability are utilized by task demands. Expertise and automatization of tasks increase capacity, while stress has a detrimental effect. Fatigue has been consistently shown to affect vigilance, attention, time estimation, and response time, along with an array of physiological indices.

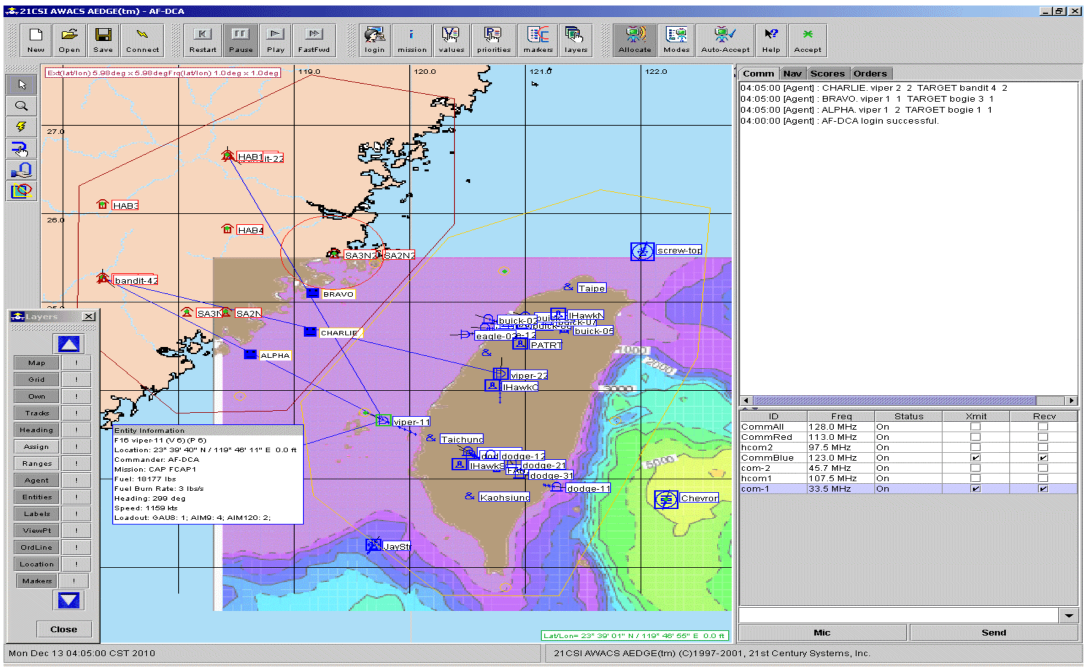
Fatigue is also expected to affect motivation, which in turn should also add to the prediction of performance, after consideration of capacity. Motivation is expected to moderate the relationship between capacity and performance. That is, the relationship between capacity and performance depends in part on the level of motivation. High capacity will not necessarily result in high performance when motivation is low. In general, increased fatigue is expected to reduce motivation and thus negatively affect shared awareness As data accumulates, hierarchical linear regression models will be used to test the model for specific interaction and mediating effects.

Individual performance is expected to mediate team decisionmaking (Hollenbeck et al., 1996) and predict general performance outcomes. Our initial data collection focused on individual performance within a team context, with initial hypotheses regarding general detrimental effect on individual performance. Subsequent studies will add to analyses, through investigations more focused on teamwork, decision support, and interventions to ameliorate or delay effects of fatigue.

Method

Data collection was conducted in the AFRL Cognitive Assessment and Sleep Laboratory (CASL) located within Bldg. 1192 at Brooks AFB or in the AFRL facility located in Bldg 170, Brooks AFB.

The experimental platform is referred to as the Agent Enabled Decision Group Environment (AEDGE). The AEDGE is constructed as a federation of intelligent agent-based functions that enable PC-based scenario construction and the emulation of C2 information and cognitive task demands . These PC-based scenarios operate much like networked videogames. However, the scenarios are driven from cognitive task analyses of a particular operational context, and enable manipulations and measures for research goals.



The AWACS AEDGE was conceived through cognitive and functional analysis of team member roles, responsibilities, and decision making (Chaiken et al., 2001), to optimize generalizability of results to operational settings. Systematic descriptions of AWACS roles, responsibilities, requirements, interdependencies, tactics, strategies, and task demands were collected from subject matter experts, cognitive task analyses (Fahey et al., 1998; MacMillan et al., 1998) and focal-group interviews (Elliott et al., 1999; Elliott et al., 2001). These data were further integrated to generate a cognitive-functional taxonomy for AWACS WDs (Coovert, 1999). These data were examined to identify teamwork and decision events which were generic to performance, regardless of mission scenario, and likely to bottleneck under high tempo situations.

The AWACS-AEDGE, built using 21st Century Systems Inc.'s AEDGE™ infrastructure, is a distributed, real-time team decision support environment comprised of simulators, entity framework, intelligent agents and user interfaces. The environment supports a wide variety of

air, sea (surface and sub-surface), and ground assets in a combat environment, primarily based on the roles and responsibilities of AWACS WD team members, but including a variety of military platforms and weapons, with realistic but unclassified capabilities. The environment has been tested with an excess of two hundred physical entities (planes, ships, SAM sites, etc.) operating with realistic yet non-classified performance characteristics in an interactive environment in which real-time decision support is available to each WD.

The behavior and decisionmaking of all hostile and friendly entities not controlled by humans is directed by agent-based technology. This results in several related capabilities. First, agents can “play” all roles not played by a human operator. This enables highly controlled investigation of individual performance within a team setting, where the expertise and performance of the other “players” can be controlled. In addition, this same capability provides optional decision support. If a human decides to “log in” as a particular entity, he/she may choose to view and accept recommendations generated by the agent for that entity. Characteristics of agent-based decisionmaking can be adjusted, such as degree of risk, target priorities, and general accuracy, to enable controlled investigations of performance within various information and decision support contexts.

The AEDGE architecture provides multiple levels of agent-based algorithms for decision support and modeling of performance. Generic resource allocation, search and optimization algorithms are a core part of the AEDGE product. Each AEDGE application can use and further extend these fundamental agent algorithms by either providing parameters and applications-specific values, functions and rules, or by combining, modifying or supplying new algorithms. The AWACS-AEDGE extends resource allocation, optimization and other algorithms with AWACS/WD-specific objective functions and constraints. For example, the AWACS weapon-target allocation algorithm, based on a generic resource-allocation with heuristic function evaluation, defines extended constraints such as Table 1.

```
( IN_RANGE(Target, Weapon) OR
INTERCEPT_TIME(WpnPlatform, Target) <
         MAX_TIME )
   AND
Pk(Target, Weapon) > DesiredPk(Target)
   AND
FUEL_TO_INTERCEPT(WpnPlatform, Target) +
FUEL_TO_BASE(InterceptPoint,
WpnPlatform) <
         WpnPlatform.currentFuel
```

Table 1. Sample allocation algorithm constraints

While the operator or experimenter has the option of “turning off” the decision support features, the agent recommendations will still be logged by the computer, providing rich data for study of decision modeling of computers and humans. AEDGE agent capabilities enable detailed approaches to measurement and modeling of individual and team workload, communication and decisionmaking. Tracking the number and type of recommendations generated by the agent at

any given time contributes toward new ways of conceptualizing and representing cognitive workload of individuals and teams. Agent-based recommendations may also serve as a standardized benchmark by which human tactics and decisions can be compared.

Finally, the AEDGE platform can operate through voice generation and speech recognition – operators can speak to the system using predefined jargon, request tasks be performed or information provided/transferred, and the agents will respond verbally to the speech-driven requests, using voice generation technology. All agent communications to each other, as well as to humans, are transcribed, logged to data output files, and are available online. This capability also extends capabilities for performance research, in terms of realism, efficiency, and automatic data logging.

Measures

Our initial measure of interest, reported in this paper, reflects general outcome-based performance of the individual within a team, or dyadic, context. Each participant played the role of USAF defensive counter air (DCA), and was assigned all fighter aircraft assets, along with tanking aircraft. The other USAF role was that of STRIKE, whose was assigned all bomber aircraft assets. The STRIKE role was played by the "agent", according to scripted and dynamic decision rules. The primary role of the DCA was to clear a path for the USAF bombers, and protect them. The bombers are not well-armed and will be destroyed if attacked by enemy assets. The overall score for DCA was thus based on the number of enemy assets destroyed (by participant's assets), the number of own assets destroyed, and also the number of STRIKE assets destroyed. This was referred to as the participant's "teamwork" score, as compared to the score based only on participant assets, wich was referred to as the participant's "taskwork" score. However, it should be noted that the teamwork score is not a measure of teamwork per se—it includes the taskwork score, and is based on outcomes, not teamwork processes.

Many other measures of performance were also collected. All action and decision events for each asset are logged into comprehensive data output files and form the basis of measures of team communication and coordination. In addition, agent-based recommendations were generated for the DCA role, played by each participant. These recommendations drive the behavior of all agent-based roles and of decision support features, when used. In this study, decision support features were turned off and not accessible to the participant. However, the software generates the recommendations even when the DSS feature is turned off. These recommendations will be compared to participant performance for each decision, thus enabling fine-grained investigations of decision-making. This is not straightforward for complex performance, as some recommendations are somewhat interchangeable. For example, if the participant has 2 assets that are equivalent in arms and fuel, the agent may recommend each to hostile targets that are comparable with regard to arms and location. In this case the participant would be consistent with the agent decision rules even if he/she switched the assignment of these two assets or targets. Efforts are underway to identify how to automate this comparison process, to some degree. Participants also provided subjective data with regard to performance goals and perceived workload.

Scenario construction is critical to elicitation and measurement of targeted aspects of performance in these contexts. For this effort, scenarios were constructed to be as equivalent as possible, with regard to workload and difficulty. Scenarios were constructed in collaboration with an AWACS WD subject matter expert, who began by constructing three scenarios with

similar event sequencing and type/timing of assets/targets. Each scenario had 4 roles: USAF defensive counter air (DCA), USAF STRIKE, hostile DCA, and hostile STRIKE. Each role, hostile and friendly, were assigned the same number and type of assets to start, across scenarios. Assets included air bases, aircraft (fighter, bomber, tanker), surface-to-air missile sites, and cruise missiles (USAF STRIKE). Each scenario had the same rules of engagement. In each scenario, additional resources and targets would appear regularly. The type and timing of these assets were constructed such that each five minute increment contained equivalent resource events, across scenarios. These events included high-tempo stretches of enemy activity in which many simultaneous intercept decisions had to be made while keeping track of refueling needs. Operators played the role of the USAF defensive counter-air. Agents play the USAF strike role, and equivalent enemy roles.

The WDs' primary (AFDCA) mission was to defend friendly air space with a secondary mission of protecting the AF Strike Force (slow moving, vulnerable bombers) as these carried out their primary mission of bombing enemy targets. An important property of all scenarios was their deterministic nature. Given an interceptor's weapons were in range of a target, the interceptor shot down the target with probability 1.0. While this is (perhaps) not an accurate portrayal of real-world armament, setting the simulation to be deterministic in this fashion allowed performance to be more a function of WD skill and not luck.

Method

Four groups of four research participants participated in a sustained operations research protocol. Each group participated for 5 sessions, occurring over a period of six weeks. Each session occurred over three days. The sessions would begin at 6pm after a normal day of work. Participants were kept awake until 11 am the next day. Subsequent sleep/wake schedules varied with each session Participants performed the task at 7pm, 1am, 2am, 430am, 730am, and 10am, during this first shift, and periodically throughout the session, during their time awake. Preliminary results revealed issues and challenges:

Only partial data collection was achieved for the first 2 groups. This was primarily due to the use of laptops for the task. Once the software was loaded on new desktop computers, data collection proceeded without any technical difficulties of any kind. This initial data, along with group 3, was used primarily to review measures and refine scenario content. However, remaining data do not provide sufficient statistical power to enable reasonable tests of significance. Here, we will focus on descriptive results, primarily drawn from group 3, as data analysis is ongoing. At the same time, the data will also be investigated using a multi-level hierarchical linear technique that may be able to manage the missing data such that all data can be pooled. This effort is led by Dr. Michael Coovert, at the University of South Florida.

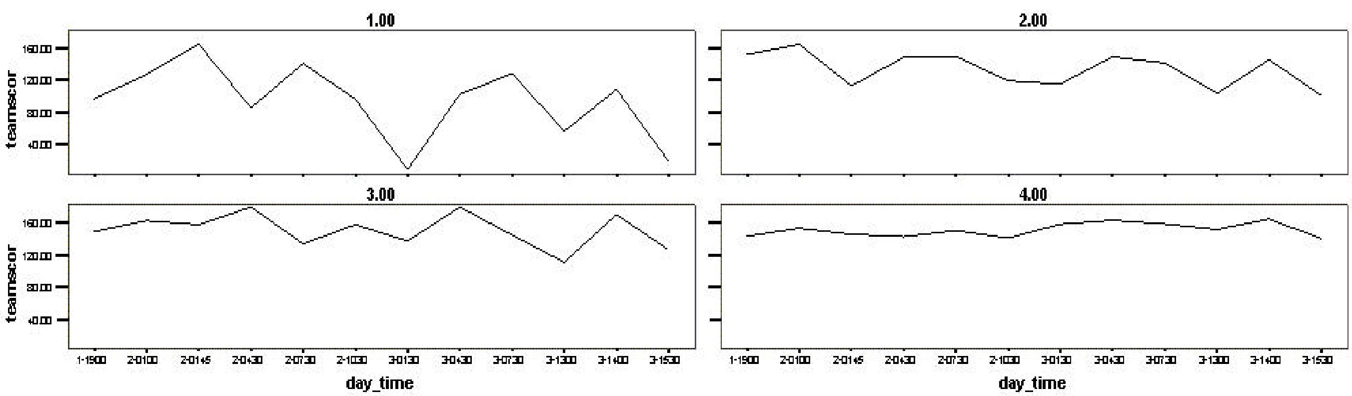
Significant differences were found due to scenario content, in unexpected ways. This was somewhat disturbing, given our systematic approach to scenario development. Half of the scenarios were developed to be more difficult; however, participant mean performance scores did not consistently follow expectations. Subsequent investigation and discussion has generated several plausible reasons for unexpected differences in scenario difficulty. For example, while assets and targets assigned to friendly and hostile roles were equivalent in type and timing of appearance, scenarios differed in the geographic distance between friendly and enemy sites—thus affecting task tempo. This was due to use of realistic maps of Taiwan, Malta, and Cyprus.

Lessons learned are currently being compiled, for use in subsequent studies dedicated to the investigation of fatigue on team performance.

There were also significant mean differences in overall performance among the participants. The differences due to scenario and individuals mask effects of fatigue, while statistical tests such as regression or ANOVA were precluded due to lack of power. Thus, even while looking at fatigue effects between two sessions using the same scenario (thus controlling for scenario difference), participants varied in performance, and perhaps also in the degree and timing of fatigue effect. Overall non-statitstical review of performance of each participant indicated that performance generally declined over time awake; however, differences due to scenario difficulty constrained any quantitative within-subjects modeling of effect over time.

Given constraints of analysis due to scenario and individual sources of variance, we then removed this additional variance, in order to 'see' remaining trends, due to fatigue or other factors. First, differences due to scenario difficulty were removed, in the same manner as in ANOVA analyses. Mean scores were calculated for each scenario and across all scenarios. The difference between scenario and overall mean was calculated for each scenario. This represents the variance due to scenario difficulty. This difference was then removed or added to each participant's score, in order to adjust the score for scenario difficulty. In the same manner, scores were further adjusted for the variance due to individual difference. Scores were adjusted based on the difference between each participant's overall mean score and the overall mean across participants. This allowed us to review scores for remaining fatigue-related effects.

Results using these adjusted scores revealed patterns of performance that were consistent with fatigue effects, in that in general, performance declined over time awake. The main finding, however, was revealed as scores were reviewed separately for each session. The charts below represent adjusted overall scores, for the last four sessions, for group 4.



Results indicate that performance improved over time, while variable diminished. Thus during the final 3-day session, all participants consistently performed well, regardless of scenario, fatigue, or time of day. This was certainly unexpected, given the complexity and task demands of the scenarios. During the last session, participants were performing consistently, at a level comparable to operational AWACS weapons directors, regardless of fatigue. We know this

as these scenarios were also used to collect performance data from weapons directors located at Tinker AFB (Chaiken et al., 2000).

As a result, participants from this group were interviewed with regard to their performance and characteristics of the scenarios. Participants agreed that while scenarios differed in difficulty, all became easier over time, as they learned general tactics and coping strategies to meet task demands. For example, they learned to move assets early in the scenario, for tactical advantage. They also learned interface strategies, such as using pull-down information tables to quickly assess armament and fuel level for each asset. Also, coordination of task demands became easier as procedures became more habitual. Participants also reported that task management became less demanding, as routine tasks, such as monitoring asset information, became more habitual.

Discussion

Results reported here are preliminary, and focus on lessons learned. Additional data analyses, focused on scenario events and decisions, are ongoing. Alternative modeling techniques will be further explored for use in team-based repeated measures context.

Results underline challenges inherent to repeated-measures testing using complex performance criteria. Fatigue research is typified by use of fewer participants for longer periods of time, and within-subjects analysis of effect. This approach requires that multiple assessments are comparable, such that trends in performance can be detected. This is not a problem when using relatively simple tests of perceptual or cognitive ability, such as processing speed or working memory.

The challenge arises when more complex scenario-based performance is of interest. A variety of scenarios is needed, to minimize practice effects. In this effort, six scenarios were systematically developed such that they would appear different but have equivalent task demand. In each scenario, each role was assigned equivalent assets, which appeared at the same time in the scenario, and faced equivalent hostile assets. Characteristics such as geographic location and direction of attack were varied in order to minimize recognition and anticipation of events. Even so, differences in performance among scenarios were found and additional issues were identified relevant to C3 scenario construction.

Preliminary results are based on outcome-based data. Outcome-based data provide limited information regarding performance, and may not be sufficiently sensitive to fatigue effects unless the manipulation is more extreme. Also, outcome-based measures do not inform what kinds of error are more likely to occur, by whom, or when. Further analyses are ongoing, to identify patterns of error based on analyses of scenario events and participant decisions and sequencing of action.

Results indicated that subjects improved performance over the five 3-day session, such that all were performing well in the last session, regardless of fatigue or initial ability. This was unexpected, as subjects were trained to a level of performance comparable to AWACS weapons directors prior to experimental data collection, in order to minimize effects of learning. In addition, the first 3-day experimental session was not included in the analysis, to further control for skill acquisition. However, participants did indeed learn something, as performance did improve, systematically, over a relatively long period of time.

The interesting aspect of these results is that participants did improve to the point that any effect due to fatigue or initial ability was not detectable—their performance became more

resistant to effects of fatigue. This raises the question—what did they learn, and why did it help? We believe that participants ultimately acquired expertise such that their performance on these tasks was similar to performance associated with experts. As a result, coordination and execution of multiple task demands was made much less effortful, and thus more resistant to effects of fatigue. While not particularly conducive to an investigation of fatigue, it does indicate the critical role of expertise in reducing, or more likely—delaying, the inevitable effects of fatigue. Certainly, what better criterion for training evaluation than the fact that operators can perform well with less effort, and better manage effects of fatigue?

This effort was preliminary, yielding baseline information and lessons learned that will improve subsequent efforts. Additional studies of fatigue and team performance are planned, using teams of 3-4 persons, to enable in-depth investigations of team communication, decisionmaking and coordination of events. In these follow-on studies, scenarios will require a higher degree of adaptive and coordinated response from the entire team, based on inclusion of unexpected and time-critical events. This study demonstrated that even complex performance can be made, to a degree, routine. However, we expect that non-routine and time-critical events, requiring cognitive and interactive processes such as problem recognition, consensus-building, and resolution, will be more vulnerable to effects of fatigue; that is to say, the inevitable effects of fatigue will be detected earlier.

References

Barnes, C., Elliott, L. R., & Entin, E. (2001). Employing Internet2 Technology to Enable Collaborative Research and Distributed Training in Complex Multi-Operator Settings. WebNet Journal, 3 (4), 24-31.

Bowers, C., Salas, E., Prince, C., & Brannick, M. (1992). Games teams play: A method for investigating team coordination and performance. Behavior Research Methods, Instruments & Computers, 24, 503-506.

Cannon-Bowers, J. A., Burns, J. J., Salas, E., & Pruitt, J. S. (1998). Advanced technology in scenario-based training. In J.A.Cannon-Bowers & E. Salas (Eds.), Making decisions under stress: Implications for individual and team training (pp. 365-374). Washington DC: APA

Chaiken, S., Elliott, L., Dalrymple, M., & Schiflett, S., Coovert, M., Riddle, D., Gordon, T., Hoffman, K., Miles, D., King, T. (2001). Weapons Director intelligent agent-assist task: Procedure and findings for a validation study. Proceedings of the International Command and Control Research and Technology Symposium, June, 2001.

Coovert, M. D., Craiger, J. P., & Cannon-Bowers, J. A. (1995). Innovations in modeling and simulating team performance: Implications for decision making. In Guzzo & E. Salas (Eds.), Team Effectivesness and decisionmaking in organizations (pp. 149-203). San Francisco: Jossey Bass.

Coovert, M. D. & Craiger, J. P. (1997). Performance modeling for training effectiveness. In J.K.Ford, S.Kozlowski, K.Kraiger, E.Salas, & M.Teachout (Eds.), Improving Training Effectiveness in Work Organizations (pp. 47-72). Mahwah, NJ: Lawrence Erlbaum Associates.

Coovert, M. D., Elliott, L., Foster, L. L., Craiger, J. P., & Riddle, D. (1999). Measurement in synthetic task environments for teams: A methodological typology. In H. J. Bullinger and J. Ziegler (Eds.), Human-Computer Interaction: Ergonomics and User Interfaces (Vol. 1. pp. 686-690). Mahwah, NJ: Lawrence Erlbaum Associates.

Eddy DR, Hursh SR (2000). Fatigue Avoidance Scheduling Tool. Phase 1 SBIR Final Report, NTI Inc., Dayton OH, to the Air Force Research Laboratory, contract F41624-99-C-6041.

Elliott, L. R., Cardenas, R., & Schiflett, S. (1999) Development and Analysis of Team-based Rating Scales Of AWACS Performance in Distributed Mission Training. Proceedings of the Annual Research and Technology Symposium on Command and Control, Newport, RI. June, 1999.

Elliott, L. R., Chaiken, S., Stoyen, A., Petrov, P., & Riddle, D. (2000). An Agent-Enabled Simulation and Training System to Support Team-based C3 Performance Research. Proceedings of the 5th Conference on Naturalistic Decision Making, Tammsvik, Sweden, 26-28 May, 2000

Elliott, L. R., Dalrymple, M., Regian, J. W., Schiflett, S. (2001). Scaling Scenarios for Synthetic Task Environments: Issues Related to Fidelity and Validity. Paper accepted for presentation at the 2001 Meeting of the Human Factors and Ergonomics Society, September.

Elliott, L. R., Hollenbeck, J. R., Schiflett, S. G., & Dalrymple, M. (2001). Investigation of Situation Awareness and Performance in Realistic Command and Control Scenarios. In M. McNeese, E. Salas, & M. Endsley (Eds.) “Group Situational Awareness: New Views of Complex System Dynamics”. HFES Press.

Fahey, R. P., Rowe, A., Dunlap, K. and DeBoom, D.(1998). Synthetic Task Design (1): Preliminary Cognitive Task Analysis of AWACS Weapons Director Teams. Technical Report. Brooks AFB, TX: Armstrong Laboratory

Hursh, S.R. (1998) Modeling Sleep and Performance within the Integrated Unit Simulation System (IUSS). Final Report for the United States Army Soldier Systems Command; Natick Research, Development and Engineering Center, Natick, Massachusetts 01760-5020; Science and Technology Directorate; Technical Report: Natick/TR-98/026L.

Kanfer, R. & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative / aptitude approach to skill acquisition. Journal of Applied Psychology, 74, 657-690.

Kanfer, R. (1991). Motivation theory and industrial and organizational psychology. In M.Dunnette (Ed.), Handbook of I/O psychology (pp. 146-170).

Kozlowski, S. W. J., Toney, R. J., Mullins, M., Weissbein, D., & Brown, K. G. (2001). Developing Adaptability: A theory for the design of integrated embedded training systems. Advances in Human Performance and Cognitive Engineering Research, 1, 59-122.

MacMillan, J, Serfaty, D., Young, P., Klinger, D., Thordsen, M., Cohen, M., & Freeman, J. (1998). A system to enhance team decision making performance: Phase 1 Final Report. Brooks AFB: Air Force Research Laboratory, Warfighter Training Research Division.

Schiflett, S. G. & Elliott, L. R. (2000). Synthetic Team Training Environments for Command and Control. In Dee Andrews and Mike Mcneese (Eds.), Aircrew Training Methods. Mahwah, NJ: Lawrence Erlbaum Associates.